PresenterPro: A tool for recording, indexing and processing prompted speech with Praat

Dellwo, Volker

Abstract: Praat (www.praat.org) is a powerful tool for a wide variety of speech analysis and processing tasks. When it comes to recording speech, however, it lacks some fundamental functions that allow a user to prompt a reader with a written list of words or sentences (henceforth: speech prompts) on a screen and index the prompted recordings for further processing. PresenterPro - a Praat plug-in - fills this gap. It (a) prompts a reader to read utterances from a screen, (b) automatically indexes the recorded speech prompts in a Praat TextGrid and (c) extracts all recorded speech prompts into individual files. It thus offers an efficient solution for recording large lists of speech prompts. The present paper describes the plug-in and discusses in which situations it is particularly useful.

PresenterPro: A tool for recording, indexing and processing prompted speech with Praat

Volker Dellwo

Department of Computational Linguistics
University of Zurich
volker.dellwo@uzh.ch

Abstract

Praat (www.praat.org) is a powerful tool for a wide variety of speech analysis and processing tasks. When it comes to recording speech, however, it lacks some fundamental functions that allow a user to prompt a reader with a written list of words or sentences (henceforth: speech prompts) on a screen and index the prompted recordings for further processing. *PresenterPro* - a Praat plug-in - fills this gap. It (a) prompts a reader to read utterances from a screen, (b) automatically indexes the recorded speech prompts in a Praat TextGrid and (c) extracts all recorded speech prompts into individual files. It thus offers an efficient solution for recording large lists of speech prompts. The present paper describes the plug-in and discusses in which situations it is particularly useful.

Index Terms: speech recordings, phonetic field work

1. Introduction

Recording speakers reading long lists of words or sentences (speech prompts) is a typical task that phoneticians or linguists need to perform frequently. Once a list of speech prompts is recorded, the post-recording editing work is usually considerable. Entire recording sessions might need to be listened to again to identify the correct version of a read prompt in case of multiple repetitions and to extract the individual utterances into files. *PresenterPro* is a Praat plug-in that makes this task easier. It presents speech prompts from a list one-by-one on a computer screen and prompts a speaker to read them. Misread items can be prompted again until produced satisfactorily. A single recording of the entire reading session is produced. Based on the time point at which a speech prompt is presented on the screen and the time point the screen is forwarded to the next prompt, a Praat TextGrid file is created in which the read versions of the speech prompts are indexed (start and end). *PresenterPro* further extracts all indexed items into individual files. It also creates a TextGrid for each extracted utterance containing the prompted text. This can be used for segment alignment with forced alignment processing, for example.

2. Why *PresenterPro*?

A variety of tools already exist which perform similar tasks compared to *PresenterPro*. Some of these tools are more sophisticated like *SpeechRecorder* (http://www.bas.uni-muenchen.de/Bas/software/speechrecorder/) or *ProRec* (http://www.phon.ucl.ac.uk/resource/prorec/) to name only two. However, *PresenterPro* hast some advantages that might not be neglected:

- **Platform independent:** *PresenterPro* runs on Praat which is a widely used speech processing and analysis tool that many phoneticians and linguists are familiar with and that is platform independent. Some good recoding tools are only available for particular platforms (e.g. ProRec only runs on Windows).
- **Free choice of recording device:** Speech recording tools typically require the use of an inbuilt recording software to carry out recordings. *PresenterPro* allows the recording to be made on any device. This makes *PresenterPro* very flexible as you can use your high-end portable recording equipment or any other preferred recording software on your computer.
- **Full back-up of entire recording session:** Some speech recording tools save the recording of a prompted utterance directly into a single file. In some cases it might be useful to be able to go back to the recording session where other helpful verbal comments or another version of read prompts can be found. *PresenterPro* leaves you with a recording of the entire recording session.
- **Easy to edit:** *PresenterPro* is written in the easily acquirable Praat scripting language. This means that changes can be applied and *PresenterPro* can be adapted to your own needs. The code does not require compilation.

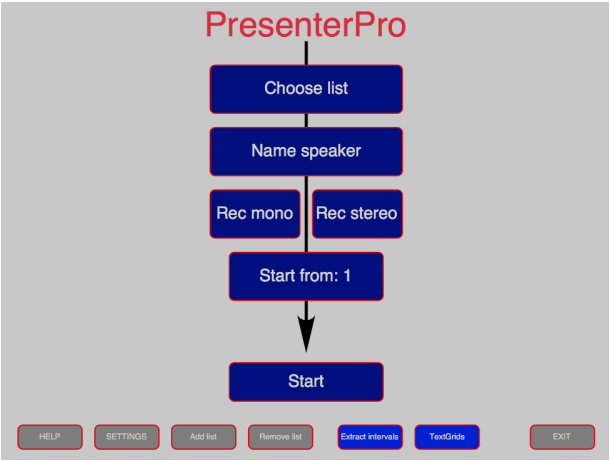

Figure 1: *Interface of PresenterPro. To carry out a recording, users work their way through the options from top to bottom along the arrow.*

3. How does *PresenterPro* work?

3.1. Obtain and install *PresenterPro*

To obtain *PresenterPro* please write an email to the author (volker.dellwo@uzh.ch). You will receive a zip directory which contains a directory named 'PresenterPro'. Use *PresenterPro* with Praat version 6.0 or higher. Install *PresenterPro* as a so called Praat plug-in. This can be done either manually by copying the directory 'plugin_PresenterPro' into your Praat preferences directory (see 'preferences directory' in the Praat help if you do not know where it is) or by executing the script 'install_deinstall_plugin.praat' that is inside the directory.

3.2. Prompting with *PresenterPro*

After installation, execute *PresenterPro* in Praat under 'New > PresenterPro'. The interface as in Fig. 1 appears in a Praat demo window. The basic idea for operating the tool is to follow the arrow along the vertical options starting under the red title 'PresenterPro' from top to bottom:

- **Choose list:** Choose a list of speech prompts from an installed set of lists. For your first usage you may want to use one of the preinstalled demo lists, for example 'demoSentences.txt'. This list contains nine speech prompts. As a side effect these prompts contain instructions about the use of PresenterPro. To add your own list of speech prompts, click on 'Add list' in the bottom menu. Your list must be a plain text file in which each speech prompt is placed in one line (do not include blank lines). To remove a list from your list selection use 'Remove list'. Lists are saved in the plug-in directory (content/lists) where they can also be added and removed manually. If you wish to see a template of a list, refer to the lists in this directory.
- **Name speaker:** Name the speaker or insert an ID. This name will be used to name the TextGrid that *PresenterPro* produces.
- **Choose your recording mode and start recording:** In case you record with Praat, click on one of the 'Rec' buttons to open the Praat Sound Recorder and start the recording inside the Sound Recorder. Choose whether you wish to record in mono or stereo mode. Stereo mode might be particularly useful if you want to record different types of signals simultaneously, like an acoustic and a laryngographic signal, for example. For single-speaker speech-only recordings, stereo is not necessary. Instead of using the Praat inbuilt sound recorder you may also use any other recording software on your computer. Instead or in addition to your computer can also start any other sound recording device at this point. All recordings will later be indexed by the Praat TextGrid that is created during the recording session.
- **Choose starting prompt:** If you do not wish to start your list from the beginning (e.g. after an interruption, see below) you can insert the number of the prompt in your list you wish to start with here.
- **Start prompting:** Start presenting your prompts by clicking on the 'Start' button. IMPORTANT: Immediately when you click the 'Start' button, *PresenterPro* will play a short calibration tone (500 Hz for 500 ms). It is very helpful to record this tone on your recording device as it will make it much easier to align your recording to the TextGrid that *PresenterPro* produces after you finish the recording. Some users reported, however, that they operate PresenterPro without the calibration tone. Bear in mind that it will make the operation more difficult.

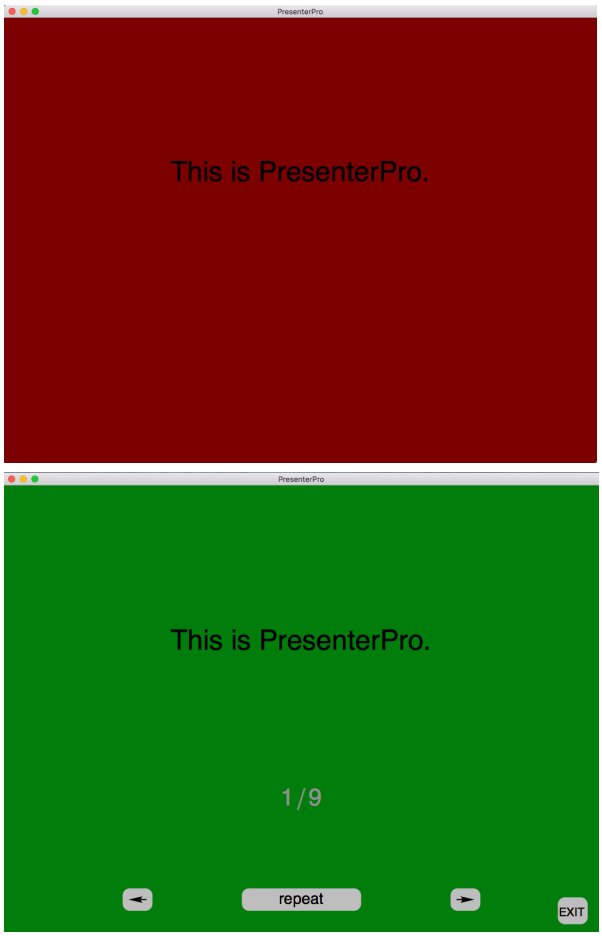

Figure 2: *A speech prompt in PresenterPro is first presented on a red screen (top) which turns to green after a short delay (bottom). The speaker's task is to read the speech prompt as soon as the screen turns green.*

- **Reading prompts:** Together with the calibration tone *PresenterPro* will also present the first speech prompt on the screen (Fig. 2). The screen will first be red and then turn green after about 1 second (delay can be changed under 'Settings'). *PresenterPro* will record the point in time at which the screen turns green and this point will be used to add a boundary in the Praat TextGrid to index the beginning of the prompted speech event (Fig. 3). It is thus essential that readers do not start reading the prompt before the screen turns green. In case the happens, press 'repeat' (see below).
- **Moving forward:** When a speaker has satisfactorily read a prompt, the forward button (arrow pointing to the right) should be pressed without much delay. This indexes the end of the reading process and places a boundary in the TextGrid respectively (Fig. 3). The next prompt will be presented immediately. If you wait for a long time before you press forward you will get long silences after you

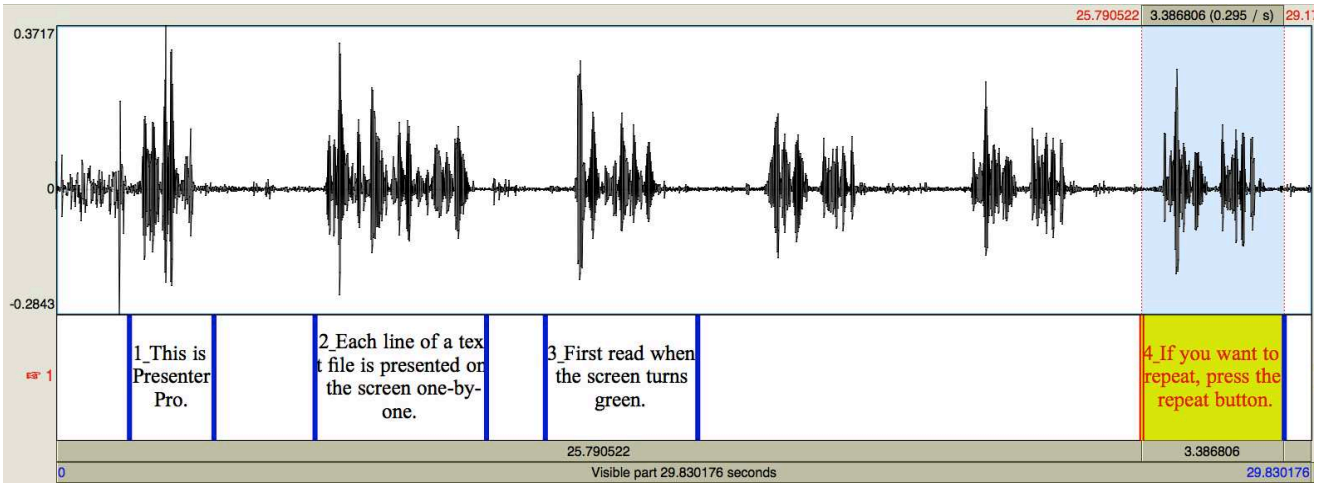

Figure 3: *Example of a recording (waveform) with a an indexed TextGrid. Filled intervals contain the prompted text. Start point of an interval is the point at which the screen turns green, end point is the point the screen is forwarded to the next prompt. Repeated speech prompts are not indexed in the TextGrid (prompt 4 was read three times in total).*

prompted utterance (or speech in case you talk). In this case just click on 'repeat' (see below) and record the speech prompt again.

- **Repeating prompts:** In case a speaker makes a mistake (e.g. reading a prompt incorrectly or reading before the screen turns green) or there is too much delay before forwarding to the next prompt, the prompt can be repeated by clicking on 'repeat'. The screen will turn red again and then green for the speaker to start reading. This process can be repeated as often as necessary. *PresenterPro* will only index the last reading of the prompt in the TextGrid. In Fig. 3 prompt number 4 was read three times in total before it was forwarded to the next prompt. Only the last version was indexed. However, all previously recorded readings of the prompt will be on the recording in case they are needed at a later point.
- **Moving backwards:** If prompting should be repeated from an earlier prompt the arrow pointing left can be used. This will delete all previously indexed boundaries up to the prompt you go back to. When moving forward, boundaries are placed again as usual.
- **Finishing prompting:** After the last prompt is presented or the 'Exit' button is pressed at any point during the prompting, a TextGrid with the intervals of the read prompts will be created and added to the Praat list. This TextGrid is also saved automatically inside the plug-in directory (content/TextGrids). When saved, the TextGrid automatically receives a time stamp. You can open any previously recorded TextGrid from this list. If you want to remove the TextGrids from that list, delete all TextGrids from your TextGrids directory inside the plug-in.
- **Calibrating your recording (with calibration tone):** Terminate your recording and load it in the Praat list of objects either from your external device or from the Praat Sound Recorder by clicking 'Add to list'. Since your recording might be rather long you may want to open it as a LongSound in Praat (see Praat help 'LongSound'). Open the recording in a Sound Editor window (View and Edit) and find the calibration tone in the beginning of the sound. Select the sound from time 0 to the end of the calibration tone as you see it in Fig. 4. Delete the selection (Edit -> Cut). Your sound file now matches your TextGrid as in Fig. 3. Save your sound under the same name as the TextGrid.
- **Calibrating your recording (without calibration tone):** In case you have not recorded a calibaration tone you need to find the starting time of your first prompted utterance in the sound and align it to the corresponding interval in your TextGrid. Measure the time of the onset of the first utterance in your Sound (utterance onset) and subtract the countdown duration from that value (by default roughly 1 sec; see settings). Measure the time of the left boundary of the corresponding interval in your TextGrid (interval onset). Subtract 'interval onset' from 'utterance onset'. The resulting value is the duration of the interval you need to remove from the start of your Sound. Make sure you backup your Sound (Save as WAV file...) in case you need several attempts for doing this (you might start to understand the value of the calibration tone at this point!).

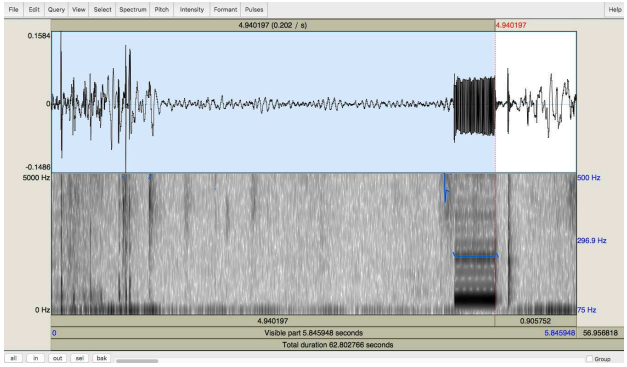

Figure 4: *Calibrating the recording by selecting the interval from time 0 to the end of the calibration tone and deleting the selection. The recording will then be aligned to the TextGrid.*

- **Extract intervals to sound files:** Once you have a Sound or LongSound with a corresponding TextGrid in your list

of objects, select both (and only these two) and click on 'Extract intervals'. Fill in the menu and continue. All your prompted utterances will now be saved to individual files. File names contain the item number. A TextGrid with the interval text will be created along with each sound file. With this TextGrid you can auto-align the segments in your files using Praat or other forced alignment tools.

Other helpful information:

- **Record it all:** After you have started your recording there is no need to rush as recording devices typically record for hours nowadays. If you use the Praat Sound Recorder for your recording, then make sure that your buffer size is big enough to allow long enough recordings (Praat > Preferences > Sound recording preferences...; see 'SoundRecorder' in Praat help menu for details).
- **Interrupting your recording sessions:** For short interruptions (e.g. speaker needs the bathroom), you can leave the recording running in *PresenterPro*. When the speaker is ready to continue you simply repeat the prompt where you left it off. For longer interruptions (e.g. overnight) you will have to stop your recording. You can exit your recording at any point with the 'Exit' button. A TextGrid will be created for the recoding up the prompt you left it off. When you restart the session you can use the 'Start from' option on the main screen (Fig. 1) and enter the number of the prompt after the one you left it off before. IMPORTANT: In case you are planning to interrupt your session you must not randomize your list of prompts (Settings>Randomize list). If you do that, both parts of your recording will contain a random selection of your list, which means that some prompts will be repeated and others will be missing.
- **Correct your alignment before extracting:** If you require start and end points that precisely align to the onset and offset of a recorded utterance you might want to manually align the boundaries in your Text Grid to your recording prior to extracting them.
- **Choose your settings:** Under 'Settings' on the main screen you can choose your line width, font size and line spacing to make sure that your sentences are presented well on the particular screen you are using (please try prior to recording). You can also choose your font and font colour. You can choose to present the prompt number and the total number of prompts under each prompt on the screen. You can also randomize your prompts for each speaker please but bear in mind that you cannot interrupt your recording anymore when you use randomization. You can also set your countdown duration here. Bear in mind that the numbers do only roughly correspond to seconds (depending on your Sound playing preferences).

4. Acknowledgements

This work was supported by the Swiss National Science Foundation (Grant number: 135287) and the Gebert Rüf Stiftung (Grant number: GRS-027/13). Many thanks to Sandra Schwab and Elisa Pellegrino for helpful comments on the draft.